

Database : Published\_Applications\_AA\_New:\*

- 1: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*
- 2: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*
- 3: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
- 4: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
- 5: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*
- 6: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*
- 7: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US11\_NEW\_PUB.pep:\*
- 8: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	434.5	33.7	261	6	US-10-975-697-6	Sequence 6, Appli
2	434.5	33.7	267	6	US-10-975-697-14	Sequence 14, Appl
3	431.5	33.4	261	6	US-10-975-697-10	Sequence 10, Appl
4	431.5	33.4	261	6	US-10-975-697-12	Sequence 12, Appl
5	430.5	33.4	261	6	US-10-975-697-8	Sequence 8, Appli
6	265.5	20.6	259	7	US-11-180-275A-24	Sequence 24, Appl
7	235	18.2	206	6	US-10-370-959-85	Sequence 85, Appl
8	225.5	17.5	158	7	US-11-249-111-85	Sequence 85, Appl
9	178.5	13.8	260	6	US-10-196-749-378	Sequence 378, App
10	178.5	13.8	260	7	US-11-101-316-128	Sequence 128, App
11	156	12.1	254	7	US-11-266-747-170	Sequence 170, App
12	156	12.1	254	7	US-11-266-747-338	Sequence 338, App
13	156	12.1	254	7	US-11-266-747-396	Sequence 396, App
14	155.5	12.1	254	7	US-11-266-747-222	Sequence 222, App
15	155.5	12.1	300	6	US-10-196-749-18	Sequence 18, Appl
16	155	12.0	254	7	US-11-266-747-402	Sequence 402, App
17	153.5	11.9	254	7	US-11-266-747-680	Sequence 680, App
18	153	11.9	254	7	US-11-266-747-278	Sequence 278, App
19	153	11.9	254	7	US-11-266-747-342	Sequence 342, App

Database : Published\_Applications\_AA\_Main:\*

- 1: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 3: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*
- 4: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*
- 5: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*
- 6: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US11\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1290	100.0	254	4	US-10-669-503-1	Sequence 1, Appli
2	944	73.2	254	3	US-09-978-758-2	Sequence 2, Appli
3	944	73.2	254	4	US-10-766-421-2	Sequence 2, Appli
4	575.5	44.6	248	5	US-10-501-282-874	Sequence 874, App
5	575.5	44.6	281	5	US-10-501-282-876	Sequence 876, App
6	556.5	43.1	246	4	US-10-369-493-237	Sequence 237, App
7	488.5	37.9	261	4	US-10-282-122A-57843	Sequence 57843, A
8	467	36.2	264	5	US-10-475-150B-4	Sequence 4, Appli
9	465.5	36.1	261	5	US-10-681-207-2	Sequence 2, Appli
10	465.5	36.1	261	5	US-10-681-207-6	Sequence 6, Appli
11	460.5	35.7	261	4	US-10-639-159-12	Sequence 12, Appl
12	460.5	35.7	261	4	US-10-782-258-12	Sequence 12, Appl
13	460.5	35.7	261	5	US-10-915-927-4	Sequence 4, Appli
14	458.5	35.5	261	4	US-10-369-493-16734	Sequence 16734, A
15	458.5	35.5	261	5	US-10-915-927-52	Sequence 52, Appl
16	457.5	35.5	261	3	US-09-971-536-74	Sequence 74, Appl

## RESULT 2

US-09-978-758-2

```

; Sequence 2, Application US/09978758
; Publication No. US20020192783A1
; GENERAL INFORMATION:
; APPLICANT: Kudoh, Masatake
; APPLICANT: Yamamoto, Hiroaki
; TITLE OF INVENTION: (R)-2-OCTANOL DEHYDROGENASES, METHODS
; TITLE OF INVENTION: FOR PRODUCING THE ENZYMES, DNA ENCODING THE ENZYMES, AND
; TITLE OF INVENTION: METHODS FOR PRODUCING ALCOHOLS USING THE ENZYMES
; FILE REFERENCE: 06501-090001
; CURRENT APPLICATION NUMBER: US/09/978,758
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: PCT/JP01/01082
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: JP 2000-374593
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: JP 2000-43506
; PRIOR FILING DATE: 2000-02-16
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Pichia finlandica
US-09-978-758-2

```

```

Query Match          73.2%; Score 944; DB 3; Length 254;
Best Local Similarity 69.7%; Pred. No. 1.8e-78;
Matches 177; Conservative 36; Mismatches 41; Indels 0; Gaps 0;

```

```

Qy      1 MSYNFANKVLIVTGGLSGIGLAVAKKFLQLGAKVTISDISATEKYNTVVGEFKTEGIDVK 60
        ||||| ||| :||| |||||:|||||:|||||:| :||: | | : ::
Db      1 MSYNFHNKVVAVVTGALSGIGLSVAKKFLQLGAKVTISDVSGEKKYHETVVALKAQNLNTD 60

Qy      61 NVQYIQADASKEADNEKLISETLSAFGLDYVCANAGIATFTQTDISYDVWRKVTSINL 120
        |: |:|:|:| | |:|:|:|:| | | | | | | | | | | : |:|:| | :|:|
Db      61 NLHYVQADSSKEEDNKKLISETLATFGGLDIVCANAGIGKFAPTHETPFVDVWKKVIAVNL 120

Qy      121 DGVFMLDKLAAQYFLSKNKPGAIVNMGSISYVAAPGLSHYGAAGGGLKLLTQTMALLEYA 180
        :|:|:|:| | |:| |:| | | | | | | | | | | | | | | | | | | |
Db      121 NGVFLDLKLAINYWLEKSKPGVIVNMGSVHSFVAAPGLAHYGAAGGVKLLTQTLALLEYA 180

Qy      181 AKGIRVNSVNPgyIKTPLLDICPKEHMDYLITQHPIGRIGKPEEIASAVAF LCSDEATFI 240
        : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      181 SHGIRVNSVNPgyISTPLIDEVPKERLDKLVSLHPIGRLGRPEEVADAVAF LCSQEATFI 240

Qy      241 NGISLLVDGGYTAR 254
        ||:| | | | | | :
Db      241 NGVSLPVDGGYTAQ 254

```

Database : A\_Geneseq\_8:\*  
 1: geneseqp1980s:\*  
 2: geneseqp1990s:\*  
 3: geneseqp2000s:\*  
 4: geneseqp2001s:\*  
 5: geneseqp2002s:\*  
 6: geneseqp2003as:\*  
 7: geneseqp2003bs:\*  
 8: geneseqp2004s:\*  
 9: geneseqp2005s:\*  
 10: geneseqp2006s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Query		Length	DB	ID	Description
		Match					
1	1290	100.0		254	4	AAB49773	Aab49773 Protein w
2	944	73.2		254	4	AAB47522	Aab47522 (R)-2-oct
3	944	73.2		254	9	AED95198	Aed95198 Propargyl
4	575.5	44.6		248	6	ADB06934	Adb06934 Alloiococ
5	575.5	44.6		281	6	ADB06936	Adb06936 Alloiococ
6	573.5	44.5		257	7	ADF06077	Adf06077 Bacterial
7	556.5	43.1		246	8	ADN17584	Adn17584 Bacterial
8	488.5	37.9		261	6	ABU29919	Abu29919 Protein e
9	467	36.2		264	6	ABP60424	Abp60424 Lactobaci

Database : Issued\_Patents\_AA:\*

- 1: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/5\_COMB.pep:\*
- 2: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/6\_COMB.pep:\*
- 3: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/7\_COMB.pep:\*
- 4: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/H\_COMB.pep:\*
- 5: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/PCTUS\_COMB.pep:\*
- 6: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/RE\_COMB.pep:\*
- 7: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	944	73.2	254	2	US-09-978-758-2	Sequence 2, Appli
2	573.5	44.5	257	2	US-09-543-681A-6362	Sequence 6362, Ap
3	517.5	40.1	309	2	US-09-248-796A-16823	Sequence 16823, A
4	456.5	35.4	261	2	US-10-004-115B-36	Sequence 36, Appl
5	443	34.3	252	2	US-08-822-322-8	Sequence 8, Appli
6	443	34.3	252	2	US-09-466-109-8	Sequence 8, Appli
7	441.5	34.2	261	2	US-09-468-738A-29	Sequence 29, Appl
8	441.5	34.2	261	2	US-09-940-019-29	Sequence 29, Appl
9	441.5	34.2	261	2	US-09-940-037A-29	Sequence 29, Appl
10	441.5	34.2	261	2	US-09-305-390-18	Sequence 18, Appl
11	441.5	34.2	261	2	US-09-855-309-29	Sequence 29, Appl
12	432.5	33.5	263	2	US-09-710-279-204	Sequence 204, App
13	430.5	33.4	267	2	US-09-134-001C-5042	Sequence 5042, Ap
14	429.5	33.3	251	2	US-09-648-004-20	Sequence 20, Appl
15	429.5	33.3	251	2	US-10-272-419-20	Sequence 20,

## RESULT 1

US-09-978-758-2

; Sequence 2, Application US/09978758

; Patent No. 6706507

; GENERAL INFORMATION:

; APPLICANT: Kudoh, Masatake

; APPLICANT: Yamamoto, Hiroaki

; TITLE OF INVENTION: (R)-2-OCTANOL DEHYDROGENASES, METHODS

; TITLE OF INVENTION: FOR PRODUCING THE ENZYMES, DNA ENCODING THE ENZYMES, AND

; TITLE OF INVENTION: METHODS FOR PRODUCING ALCOHOLS USING THE ENZYMES

; FILE REFERENCE: 06501-090001

; CURRENT APPLICATION NUMBER: US/09/978,758

; CURRENT FILING DATE: 2002-04-19

; PRIOR APPLICATION NUMBER: PCT/JP01/01082

; PRIOR FILING DATE: 2001-02-15

; PRIOR APPLICATION NUMBER: JP 2000-374593

; PRIOR FILING DATE: 2000-12-08

; PRIOR APPLICATION NUMBER: JP 2000-43506

; PRIOR FILING DATE: 2000-02-16

; NUMBER OF SEQ ID NOS: 20

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 254

; TYPE: PRT

; ORGANISM: Pichia finlandica

US-09-978-758-2

Query Match 73.2%; Score 944; DB 2; Length 254;

Best Local Similarity 69.7%; Pred. No. 4.6e-95;

Matches 177; Conservative 36; Mismatches 41; Indels 0; Gaps 0;

```

Qy      1 MSYNFANKVLIVTGGLSGIGLAVAKKFLQLGAKVTISDISATEKYNTVVGEFKTEGIDVK 60
        ||||| ||| :||| |||||:|||||:|||||:| :||: | : ::
Db      1 MSYNFHNKVAVVTGALSGIGLSVAKKFLQLGAKVTISDVSGEKKYHETVVALKAQNLNTD 60

Qy     61 NVQYIQADASKEADNEKLISETLSAFGLDYVCANAGIATFTQTDISYDVWRKVTSINL 120
        |: |:|:|:| | |:|:|:|:| | | | | | | | | | | : |:|:| |:|
Db     61 NLHYVQADSSKEEDNKKLISETLATFGGLDIVCANAGIGKFAPTHETPFVDVWKKVIAVNL 120

Qy    121 DGVFMLDKLAAQYFLSKNKPGAIVNMGSISYVAAPGLSHYGAAGGGLKLLTQTMALLEYA 180
        :|:|:|:| | |:| |:| | | | | | | | | | | | | | | | | | |
Db    121 NGVFLDLKLAINYWLEKSKPGVIVNMGSVHSFVAAPGLAHYGAAGGVKLLTQTLALEYA 180

Qy    181 AKGIRVNSVNPGYIKTPLLDICPKEHMDYLITQHPIGRLGKPEEIASAVAF LCSDEATFI 240
        : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    181 SHGIRVNSVNPGYISTPLIDEVPKERLDKLVSLHPIGRLGKPEEVADAVAF LCSQEATFI 240

Qy    241 NGISLLVDGGYTAR 254
        ||:| | | | | | :
Db    241 NGVSLPVDGGYTAQ 254

```

## RESULT 2

US-09-543-681A-6362

; Sequence 6362, Application US/09543681A

; Patent No. 6605709

; GENERAL INFORMATION:

; APPLICANT: GARY BRETON

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRAB

; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.1002-001

; CURRENT APPLICATION NUMBER: US/09/543,681A

; CURRENT FILING DATE: 2000-04-05

; PRIOR APPLICATION NUMBER: US 60/128,706

; PRIOR FILING DATE: 1999-04-09

; NUMBER OF SEQ ID NOS: 8344

; SEQ ID NO 6362

; LENGTH: 257

; TYPE: PRT

; ORGANISM: Proteus mirabilis

US-09-543-681A-6362

Query Match 44.5%; Score 573.5; DB 2; Length 257;

Best Local Similarity 45.8%; Pred. No. 2.3e-54;

Matches 119; Conservative 47; Mismatches 69; Indels 25; Gaps 4;

```

Qy      4 NFANKVLIVTGGLSGIGLAVAKKFLQLGAKVTISDISATEKYNTVVGEFKTEGIDVKN-- 61
      :|:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      13 SFSNKVGIIITGGNSGIGLASAEKFLSLGASVVIT-----GSHSGRGEKAQNYL 60

Qy      62 -----VQYIQADASKEADNEKLISETLSAFGDLDYVCANAGIATFTQTTDISYDVWRKV 115
      ::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      61 QDKGFSAEFVQMDAANEQDNQQLVKYVVKHYGKVDFLFANAGVLTDTVADKLHYDQWKKV 120

Qy      116 TSINLDGVFMLDKLAAQYFLSKNKP GAI VNMGSIHSYVAAPGLSHYGAAKGGLKLLTQTM 175
      :|:|:|:|:|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      121 LDVNLNGLFLINRAVIQHWLETKTGTGAIVNCSSICSFVGQHAFPAYCSSKGGIKLLTQTL 180

Qy      176 ALEYAAKGIRVNSVNP GYIKTPLL DICPKEHMD--YLITQHPIGR LGKPEEIASAVAF LC 233
      | | : | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      181 ALDYASQGIRVNAVCPGYIDTPLLE---GRELDKQKLVALHPIGRLGTPEEVANVVAFLA 237

Qy      234 SDEATFINGISLLVDGGGYTA 253
      | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      238 SDAASFVTGASYLVDGGFTA 257

```

Database : PIR\_80:\*  
 1: pir1:\*  
 2: pir2:\*  
 3: pir3:\*  
 4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	% Query		Length	DB	ID	Description
	Score	Match				
1	466.5	36.2	261	2	I40225	glucose 1-dehydrog
2	465.5	36.1	261	1	S00812	glucose 1-dehydrog
3	460.5	35.7	261	2	I40224	glucose 1-dehydrog
4	457.5	35.5	261	2	JS0385	glucose 1-dehydrog
5	457.5	35.5	261	2	I39853	glucose 1-dehydrog
6	451.5	35.0	261	2	A33528	glucose 1-dehydrog
7	441.5	34.2	261	2	D69629	glucose 1-dehydrog
8	437	33.9	262	2	S02299	glucose 1-dehydrog
9	434.5	33.7	263	2	S01227	glucose 1-dehydrog
10	418	32.4	272	2	AG3359	glucose 1-dehydrog
11	410.5	31.8	255	1	S39737	glucose 1-dehydrog
12	407.5	31.6	245	2	AE3517	3-oxoacyl-(acyl-ca
13	393.5	30.5	299	2	D90481	hypothetical prote
14	390.5	30.3	269	2	AE2285	glucose 1-dehydrog
15	386.5	30.0	258	2	G69755	glucose 1-dehydrog



Database : UniProt\_7.2:\*  
 1: uniprot\_sprot:\*  
 2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	605.5	46.9	247	2	Q81AU6_BACCR	Q81au6 bacillus ce
2	550	42.6	246	2	Q8G0H2_BRUSU	Q8g0h2 brucella su
3	522.5	40.5	262	2	Q5A1C0_CANAL	Q5a1c0 candida alb
4	513	39.8	253	2	Q46PX5_RALEJ	Q46px5 ralstonia e
5	494	38.3	261	2	Q41AP4_9BACI	Q41ap4 exiguobacte
6	489.5	37.9	261	2	Q3XZ55_ENTFC	Q3xz55 enterococcu
7	489.5	37.9	262	2	Q5FK32_LACAC	Q5fk32 lactobacill
8	482	37.4	244	2	Q8ELC2_OCEIH	Q8elc2 oceanobacil
9	474	36.7	252	2	Q6WVP7_9LACO	Q6wvp7 lactobacill
10	466.5	36.2	261	1	DHG4_BACME	P39485 bacillus me
11	465.5	36.1	261	1	DHGA_BACME	P10528 bacillus me
12	461.5	35.8	261	2	Q9F2A6_BACLI	Q9f2a6 bacillus li
13	461.5	35.8	261	2	Q65HM4_BACLD	Q65hm4 bacillus li
14	460.5	35.7	261	1	DHG3_BACME	P39484 bacillus me
15	460.5	35.7	261	2	Q3EU76_BACTI	Q3eu76 bacillus th
16	457.5	35.5	261	1	DHG1_BACME	P39482 bacillus me